



the globus alliance  
www.globus.org

# Globus: Building Communities

Jennifer M. Schopf

Argonne National Lab

<http://www.mcs.anl.gov/~jms/Talks/>



## Agenda Today

- 2:00 - Introduction to Globus, current experience, and near term plans
  - Jennifer Schopf, ANL
- 2:30 - OGSA-DAI
  - Amy Krause, EPCC
- 2:50 – GridWay
  - Ruben Montera, Madrid
- 3:10 - Open Discussion Time
- 3:30 - end

## What is a Grid?

- Resource sharing
  - Computers, storage, sensors, networks, ...
- Coordinated problem solving
- Dynamic, multi-institutional virtual orgs
  - Community overlays on classic org structures
  - Large or small, static or dynamic
- Why is this hard?
  - Lack of central control
  - Shared resources
  - Communication and coordination

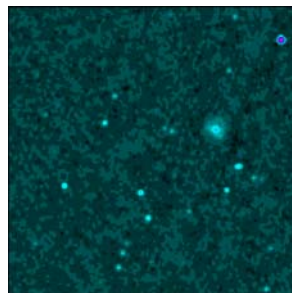
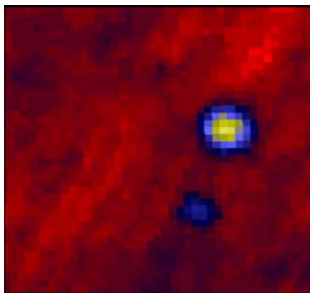
## So Why Do It?

- Computations that need to be done with a time limit
- Data that can't fit on one site
- Data owned by multiple sites
- Applications that need to be run bigger, faster, more
- Size and/or complexity of the problem requires that people in several organizations collaborate and share computing resources, data, instruments



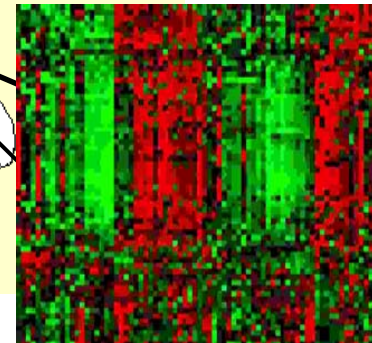
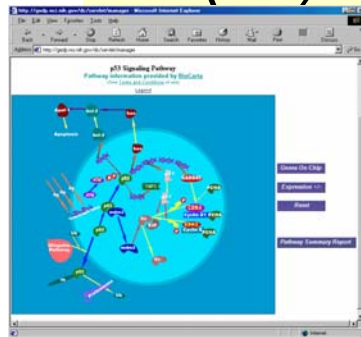
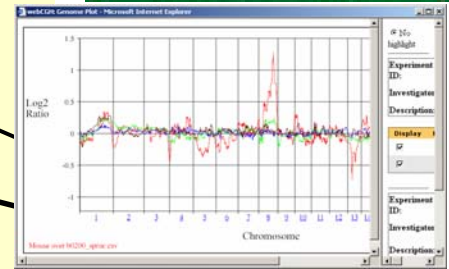
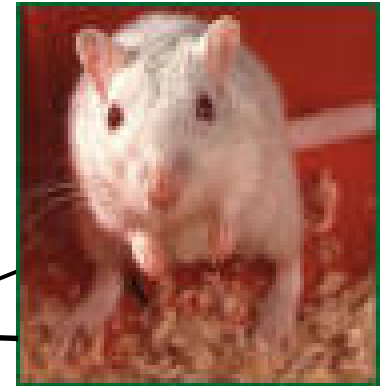
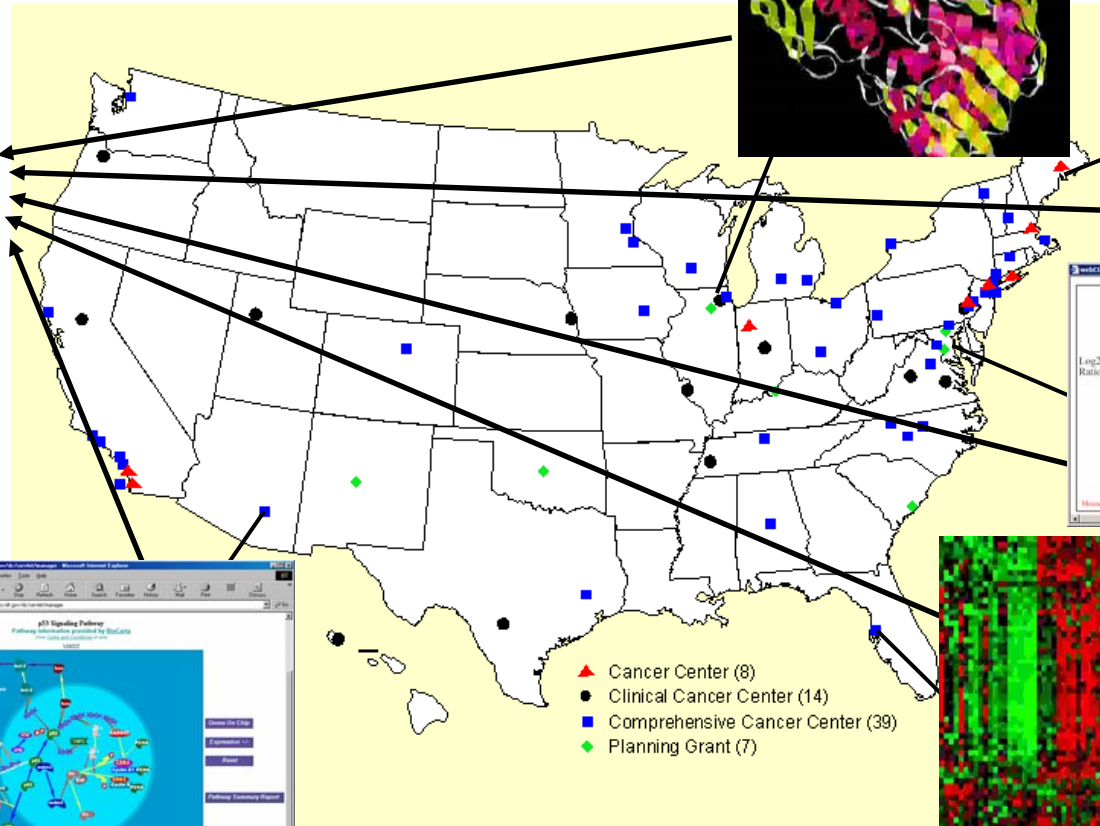
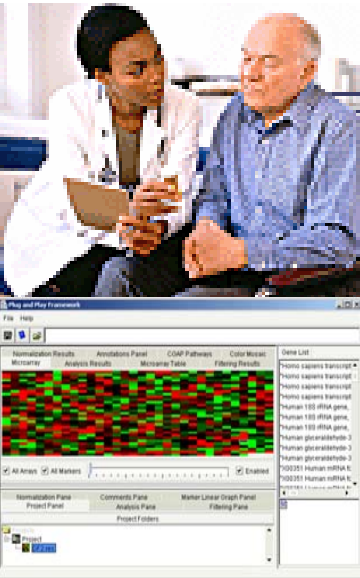
## For Example: Digital Astronomy

- Digital observatories provide online archives of data at different wavelengths



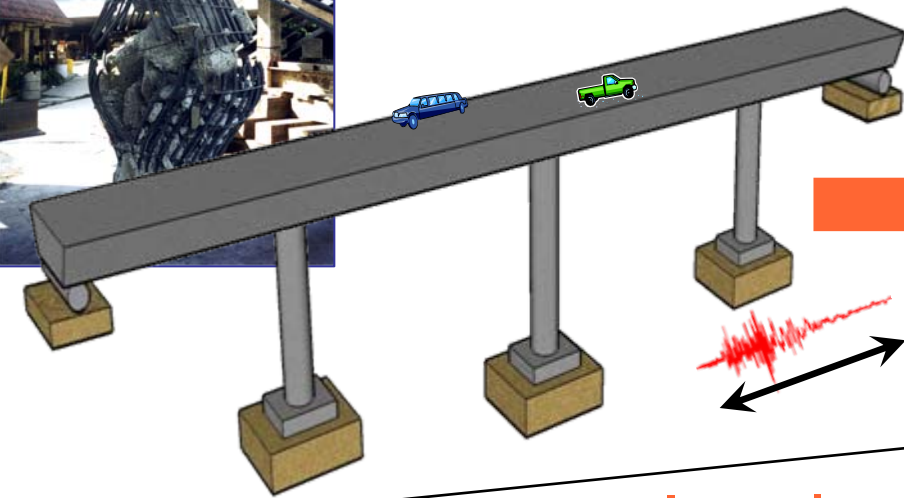
- Ask questions such as: what objects are visible in infrared but not visible spectrum?

# For Example: Cancer Biology

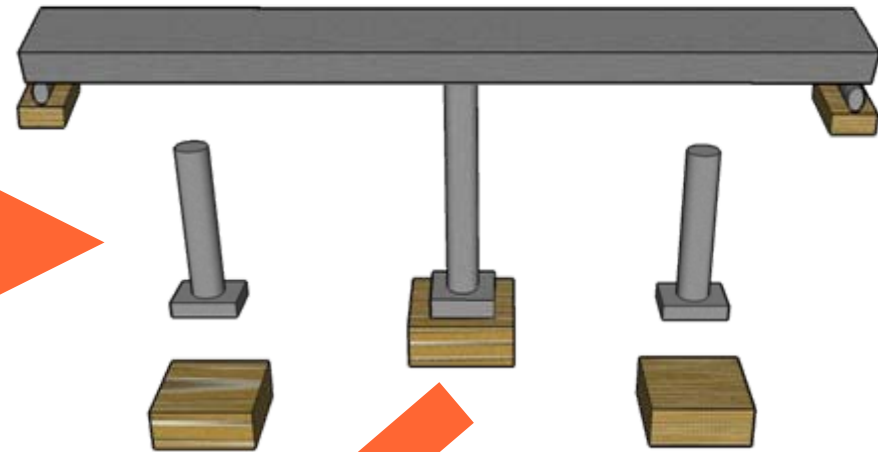


us alliance  
www.gloobus.org

## System-Level Problem

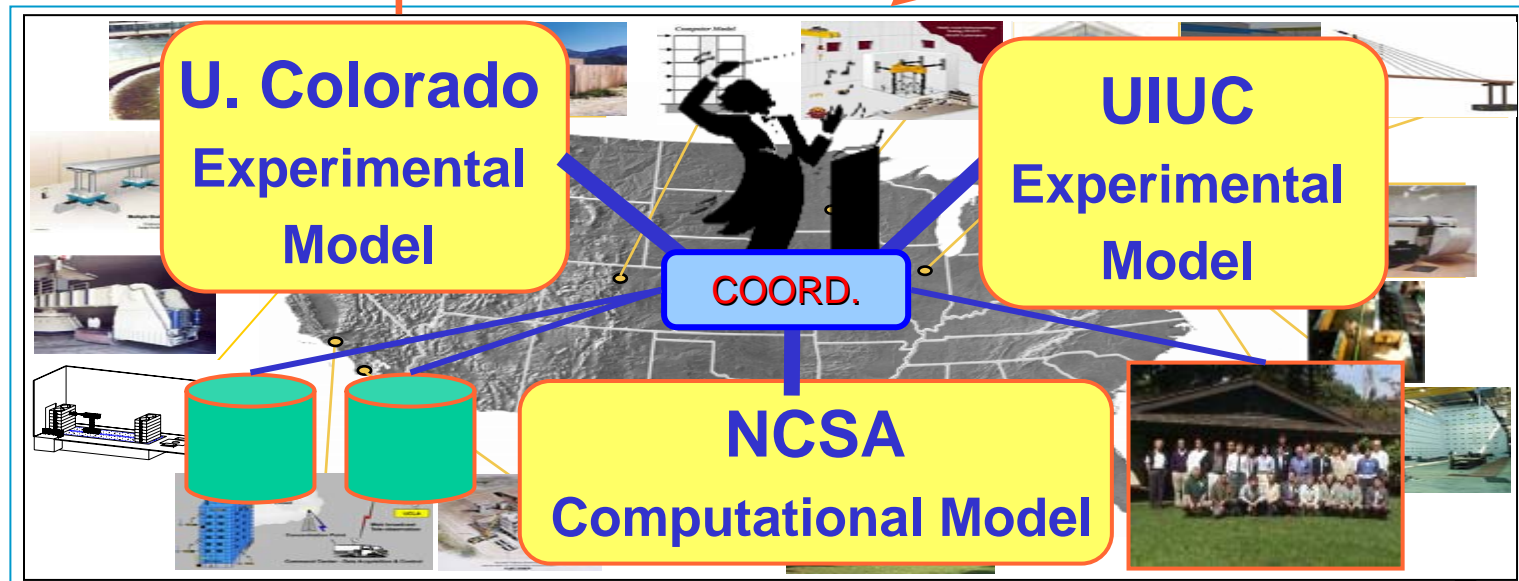


## Decomposition



## Implementation

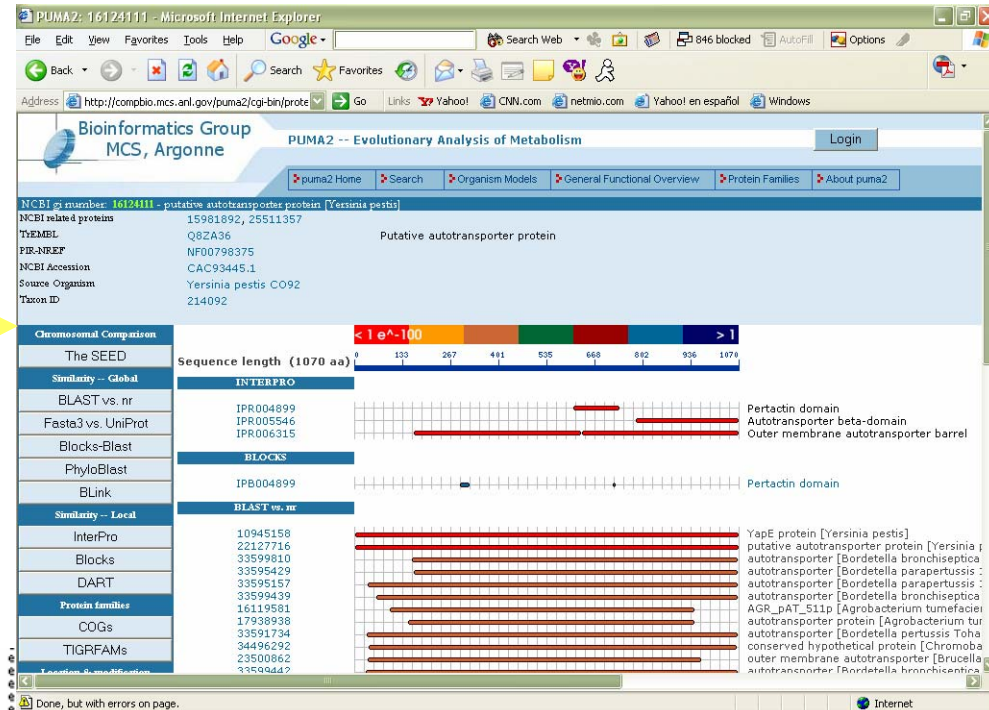
Facilities  
Computers  
Storage  
Networks  
Services  
Software  
People





# For Example: Bioinformatics

**Public PUMA  
Knowledge Base**  
Information about  
proteins analyzed  
against ~2 million  
gene sequences

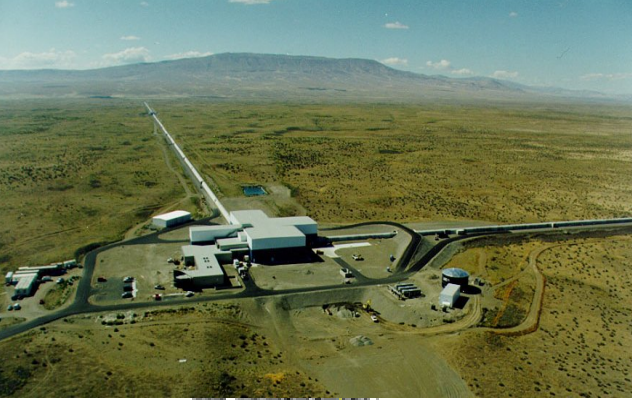


**Back Office  
Analysis**

Millions of BLAST,  
BLOCKS, etc., on  
OSG and TeraGrid

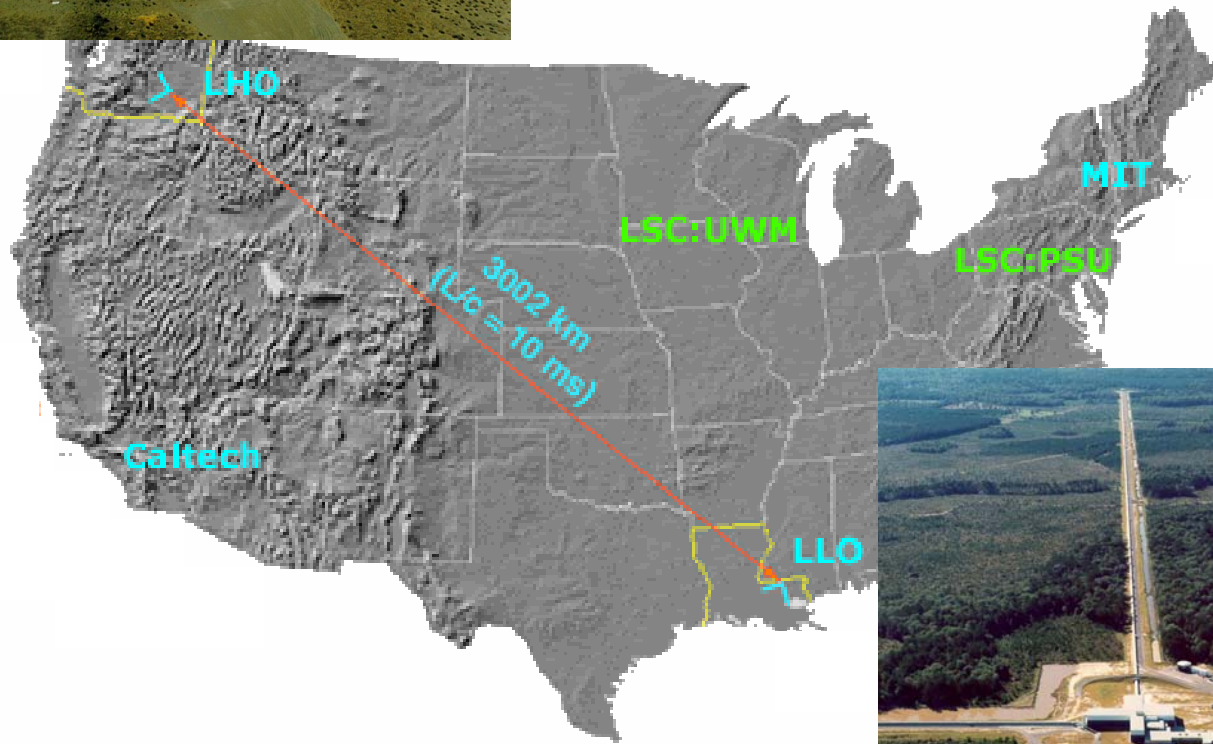
|                                 |                                 |       |     |     |   |    |     |    |        |   |           |
|---------------------------------|---------------------------------|-------|-----|-----|---|----|-----|----|--------|---|-----------|
| gi 23469780 gn REF_tigr BRA0013 | gi 16080253 ref NP_391080.1     | 44.27 | 253 | 131 | 1 | 15 | 257 | 8  | 2603.7 | e | -30.134.4 |
| gi 23469780 gn REF_tigr BRA0013 | gi 23088409 ref NP_691875.1     | 43.46 | 253 | 133 | 2 | 16 | 258 | 5  | 2573.8 | e | -30.133.7 |
| gi 23469780 gn REF_tigr BRA0013 | gi 48837187 ref ZP_00294182.1   | 44.92 | 256 | 125 | 2 | 14 | 256 | 7  | 2591.1 | e | -30.134.4 |
| gi 23469780 gn REF_tigr BRA0013 | gi 15205400 gb AA05342.1        | 44.75 | 257 | 136 | 2 | 15 | 258 | 3  | 2561.9 | e | -30.134.4 |
| gi 23469780 gn REF_tigr BRA0013 | gi 48864015 ref ZP_00317908.1   | 44.49 | 245 | 134 | 1 | 13 | 257 | 5  | 2475.1 | e | -30.134.4 |
| gi 23469780 gn REF_tigr BRA0013 | gi 30348891 gb AA028934.1       | 39.53 | 253 | 138 | 3 | 18 | 257 | 5  | 2552.0 | e | -30.134.4 |
| gi 23469780 gn REF_tigr BRA0013 | gi 19655222 gb AA093939.1       | 40.64 | 251 | 138 | 1 | 17 | 256 | 10 | 2602.7 | e | -30.134.4 |
| gi 23469780 gn REF_tigr BRA0013 | gi 27358806 gb AA007757.1       | 43.03 | 251 | 130 | 4 | 18 | 256 | 11 | 2602.5 | e | -30.134.4 |
| gi 23469780 gn REF_tigr BRA0013 | gi 112597924 gb AA016899.2      | 46.70 | 182 | 96  | 1 | 62 | 243 | 5  | 1856.8 | e | -30.134.4 |
| gi 23469780 gn REF_tigr BRA0013 | gi 46363318 ref ZP_00286079.1   | 39.58 | 240 | 136 | 2 | 14 | 243 | 6  | 2361.8 | e | -30.134.4 |
| REF_tigr BRA0013                | gi 39933731 ref NP_946007.1     | 34.90 | 255 |     |   |    |     |    |        |   | -30.141.4 |
| REF_tigr BRA0013                | gi 48782600 ref ZP_00279106.1   | 35.92 | 245 |     |   |    |     |    |        |   | -30.140.2 |
| REF_tigr BRA0013                | gi 41407534 ref NP_960370.1     | 36.09 | 266 |     |   |    |     |    |        |   | -30.139.4 |
| REF_tigr BRA0013                | gi 48851585 ref ZP_00305793.1   | 32.39 | 247 |     |   |    |     |    |        |   | -30.138.0 |
| REF_tigr BRA0013                | gi 15966306 ref NP_386659.1     | 36.50 | 263 |     |   |    |     |    |        |   | -31.137.9 |
| REF_tigr BRA0013                | gi 17548526 ref NP_521866.1     | 36.36 | 264 |     |   |    |     |    |        |   | -31.137.9 |
| gi 23469780 gn REF_tigr BRA0013 | gi 151891730 ref VP_074421.1    | 38.87 | 247 | 136 | 7 | 18 | 256 | 1  | 2403.4 | e | -31.137.1 |
| gi 23469780 gn REF_tigr BRA0013 | gi 145881 gb AA023739.1         | 33.87 | 248 | 147 | 3 | 13 | 253 | 3  | 2404.4 | e | -31.137.2 |
| gi 23469780 gn REF_tigr BRA0013 | gi 125029334 ref NP_379386.1    | 35.20 | 250 | 147 | 4 | 15 | 256 | 6  | 2485.7 | e | -31.132.9 |
| gi 23469780 gn REF_tigr BRA0013 | gi 121208953 ref NP_526732.1    | 38.52 | 257 | 138 | 6 | 12 | 255 | 5  | 2545.7 | e | -31.132.9 |
| gi 23469780 gn REF_tigr BRA0013 | gi 46314028 ref ZP_00244616.1   | 35.95 | 254 | 153 | 2 | 12 | 258 | 3  | 2485.7 | e | -31.132.9 |
| gi 23469780 gn REF_tigr BRA0013 | gi 141406512 ref NP_319668.1    | 33.61 | 238 | 149 | 2 | 16 | 253 | 2  | 2309.8 | e | -31.132.1 |
| gi 23469780 gn REF_tigr BRA0013 | gi 15644471 ref NP_229523.1     | 35.69 | 255 | 144 | 5 | 12 | 256 | 2  | 2469.8 | e | -31.132.1 |
| gi 23469780 gn REF_tigr BRA0013 | gi 123470090 ref ZP_00125423.1  | 35.20 | 250 | 145 | 4 | 12 | 253 | 3  | 2439.8 | e | -31.132.1 |
| gi 23469780 gn REF_tigr BRA0013 | gi 24935279 gb AA064237.1       | 34.63 | 257 | 146 | 4 | 12 | 257 | 4  | 2499.8 | e | -31.132.1 |
| gi 23469780 gn REF_tigr BRA0013 | gi 48847665 ref ZP_00303115.1   | 36.05 | 258 | 145 | 9 | 12 | 257 | 4  | 2531.3 | e | -29.131.7 |
| gi 23469780 gn REF_tigr BRA0013 | gi 28951510 gb AA054587.1       | 36.40 | 250 | 143 | 4 | 12 | 253 | 3  | 2431.3 | e | -29.131.7 |
| gi 23469780 gn REF_tigr BRA0013 | gi 27378783 ref NP_770312.1     | 36.25 | 251 | 143 | 3 | 14 | 255 | 7  | 2491.3 | e | -29.131.7 |
| gi 23469780 gn REF_tigr BRA0013 | gi 1708836 sp P50198 LINX_PSEPA | 34.23 | 260 | 143 | 4 | 12 | 257 | 4  | 2491.7 | e | -29.131.3 |
| gi 23469780 gn REF_tigr BRA0013 | gi 33594146 ref NP_881792.1     | 34.17 | 240 | 148 | 5 | 18 | 256 | 6  | 2363.7 | e | -29.130.2 |
| gi 23469780 gn REF_tigr BRA0013 | gi 33598116 ref NP_885759.1     | 34.17 | 240 | 148 | 5 | 18 | 256 | 6  | 2363.7 | e | -29.130.2 |
| gi 23469780 gn REF_tigr BRA0013 | gi 157883206 ref NP_337874.1    | 34.20 | 241 | 148 | 5 | 18 | 256 | 6  | 2363.7 | e | -29.130.2 |





# Reliable Wide Area Data Replication

LIGO Gravitational Wave Observatory



Replicating  $>1$  Terabyte/day to 8 sites  
>30 million replicas so far

MTBF = 1 month [www.globus.org/solutions](http://www.globus.org/solutions)



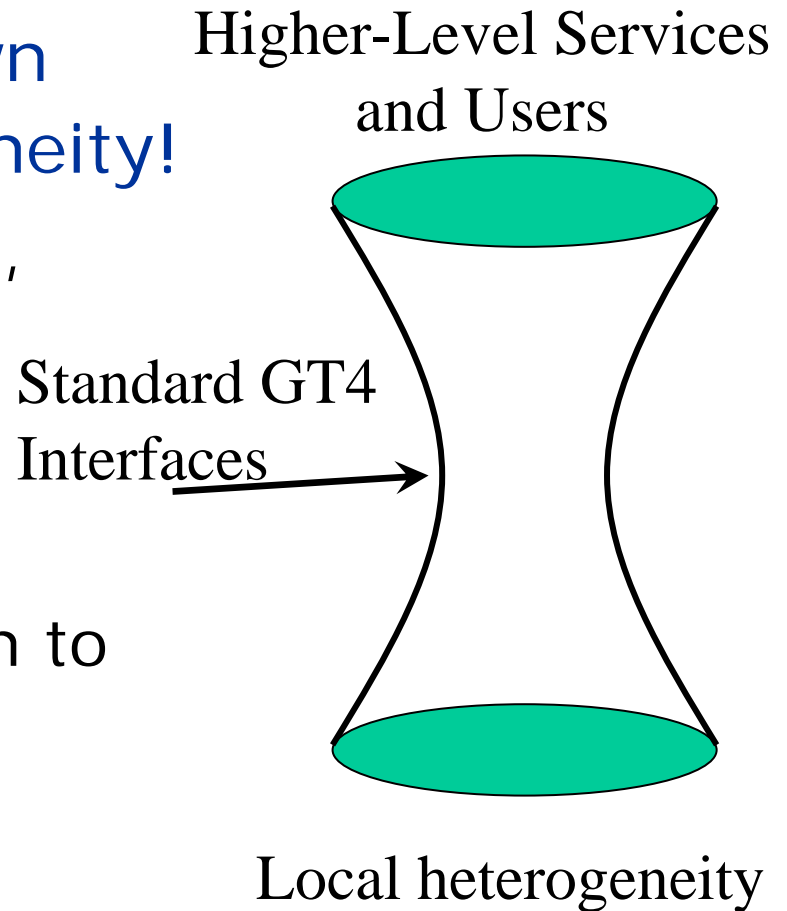
## The Role of the Globus

- A collection of solutions to problems that come up frequently when building collaborative distributed applications
- Heterogeneity
  - A focus, in particular, on overcoming heterogeneity for application developers
- Standards
  - We capitalize on and encourage use of existing standards (IETF, W3C, OASIS, GGF)
  - GT also includes reference implementations of new/proposed standards in these organizations



## Globus is an Hour Glass

- Local sites have an their own policies, installs – heterogeneity!
  - Queuing systems, monitors, network protocols, etc
- Globus unifies
  - Build on Web services
  - Use WS-RF, WS-Notification to represent/access state
  - Common management abstractions & interfaces





## Globus and dev.globus

- Globus software is organized as several dozen “Globus Projects”
  - Projects release products
- Each project has its own “Committers”
  - Committers are responsible for governance on matters relating to their products
- A “Globus Management Committee”
  - provides overall guidance and conflict resolution
  - approves the creation of new Globus projects

Guidelines  
(Apache  
Jakarta)

- [Welcome](#)
- [List of projects](#)
- [Guidelines](#)
- [Infrastructure](#)
- [How to contribute](#)
- [GlobDev events](#)
- [Recent changes](#)
- [GlobDev FAQ](#)

Infrastructure  
(CVS, email,  
bugzilla, Wiki)

#### *common runtime projects*

- [C Core Utilities](#)
- [C WS Core](#)
- [CoG jglobus](#)
- [Core WS Schema](#)
- [Java WS Core](#)
- [Python Core](#)
- [XIO](#)

#### *data projects*

- [GridFTP](#)
- [OGSA-DAI](#)
- [Reliable File Transfer](#)
- [Replica Location](#)

#### *execution projects*

- [GRAM](#)

#### *information projects*

- [MDS4](#)

#### *security projects*

- [C Security](#)
- [CAS/SAML Utilities](#)
- [Delegation Service](#)

Projects  
Include

...

[Foster](#) [my talk](#) [preferences](#) [my watchlist](#) [my contributions](#) [log out](#)  
[article](#) [discussion](#) [edit](#) [history](#) [move](#) [unwatch](#)

## Welcome

This is the new home Globus software development; it is still under construction. The current status of our efforts to build this environment can be found [on this page](#). Comments regarding this site can be sent to [info@globus.org](mailto:info@globus.org). Thank you for your interest in Globus development!

Globus was first established as an open source software project in 1996. Since that time, the Globus development team has expanded from a few individuals to a distributed, international community. In response to this growth, the Globus community (the "Globus Alliance") established in October 2005 a new source code development *infrastructure* and meritocratic *governance model*, which together make the process by which a developer joins the Globus community both easier and more transparent.

The Globus governance model and infrastructure are based on those of [Apache Jakarta](#). In brief, the governance model places control over each individual software component ([project](#)) in the hands of its most active and respected [contributors](#) (*committers*), with a [Globus Management Committee](#) (GMC) providing overall guidance and conflict resolution. The infrastructure comprises [repositories](#), [email lists](#), Wikis, and [bug trackers](#) configured to support per-project community access and management.

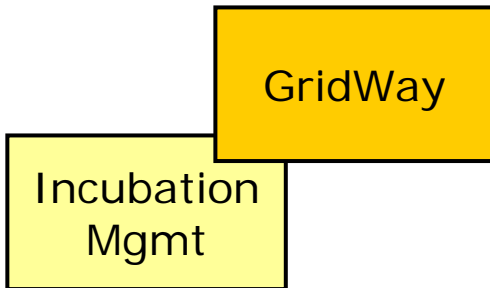
For more information, see:

- The [Globus Alliance Guidelines](#), which address various aspects of the Globus governance model and the Globus community.
- A description of the Globus Alliance [Infrastructure](#).
- A list of current Globus projects.
- Information about Globus community events.
- The [conventions and guidelines](#) that apply to contributions

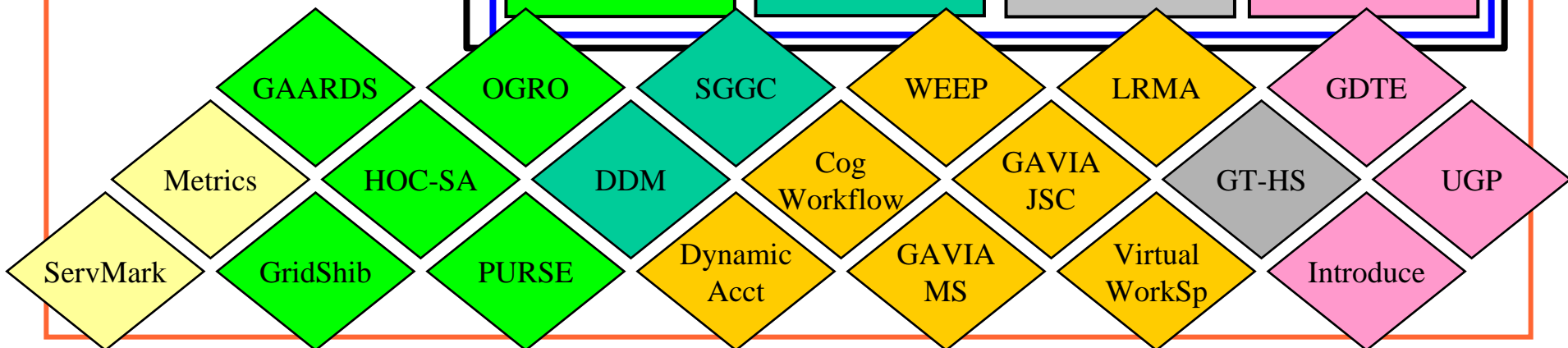
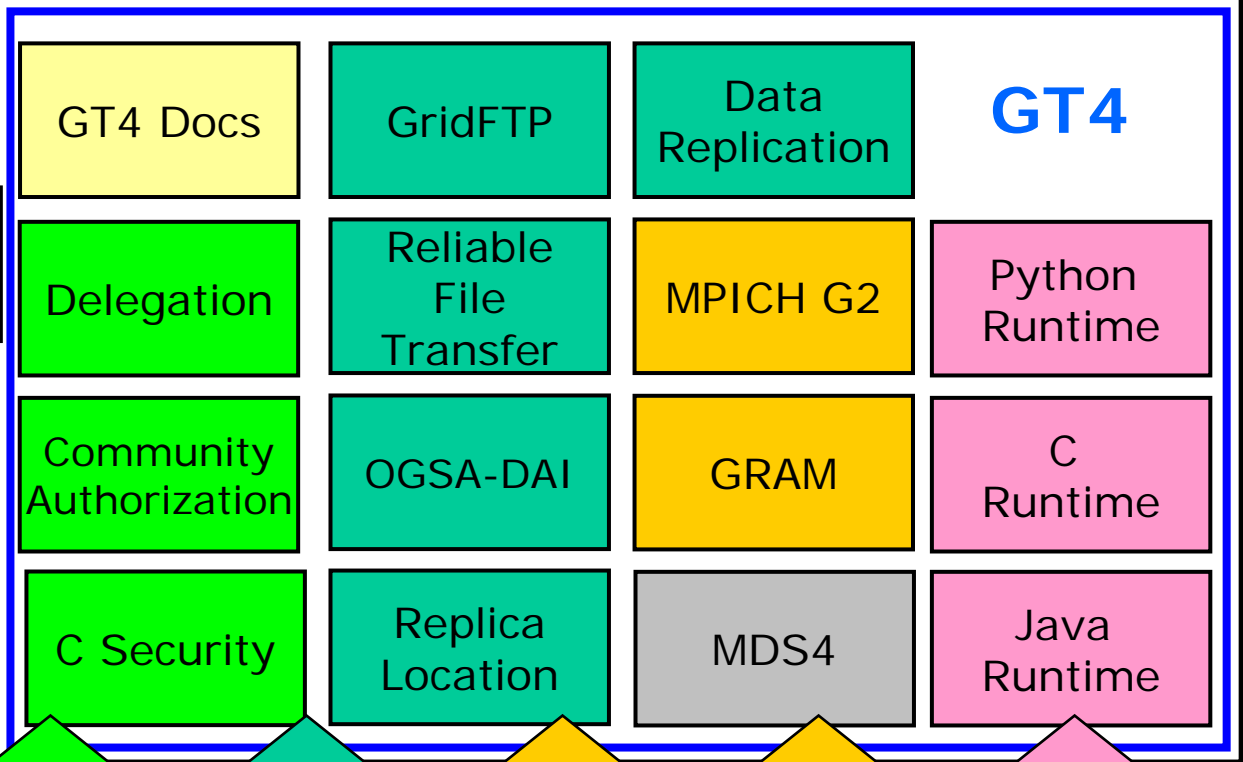


# Globus Software: [dev.globus.org](http://dev.globus.org)

## Globus Projects



## Incubator Projects







# Globus Software: [dev.globus.org](http://dev.globus.org)

## Globus Projects

GridWay

Incubation  
Mgmt

## Incubator Projects

GT4 Docs

GridFTP

Data  
Replication

**GT4**

Delegation

Reliable  
File  
Transfer

MPICH G2

Python  
Runtime

Community  
Authorization

OGSA-DAI

GRAM

C  
Runtime

C Security

Replica  
Location

MDS4

Java  
Runtime

GAARDS

OGRO

SGGC

WEEP

LRMA

GDTE

Metrics

HOC-SA

DDM

Cog  
Workflow

GAVIA  
JSC

GT-HS

UGP

ServMark

GridShib

PURSE

Dynamic  
Acct

GAVIA  
MS

Virtual  
WorkSp

Introduce

Other

Security

Data Mgmt

Execution  
Mgmt

Info  
Services

Common  
Runtime



## GT4 Core Functionality

Reference implementation of WSRF and WS-N functions

- Naming and bindings (basis for virtualization)
  - Every resource can be uniquely referenced and has one or more associated services for interacting
- Lifecycle (basis for resilient state management)
  - Resources created by svcs following a factory pattern
  - Resource destroyed immediately or scheduled
- Information model (basis for monitoring & discovery)
  - Resource properties associated with resources
  - Operations for querying and setting this info
  - Asynchronous notification of changes to properties
- Service groups (basis for registries & collective svcs)
  - Group membership rules and membership management
- Base fault type



## WSRF vs XML/SOAP

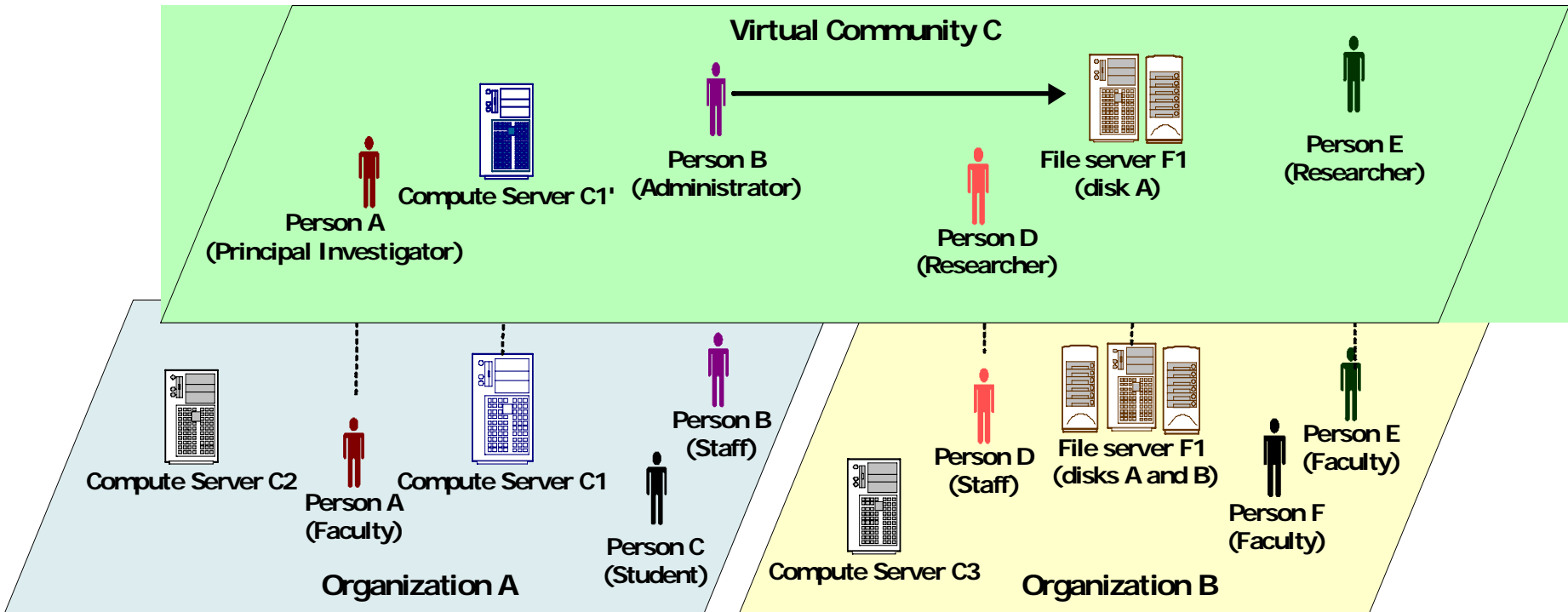
- The definition of WSRF means that the Grid and Web services communities can move forward on a common base
- Why Not Just Use XML/SOAP?
  - WSRF and WS-N *are* just XML and SOAP
  - WSRF and WS-N are just Web services
- Benefits of following the specs:
  - These patterns represent best practices that have been learned in many Grid applications
  - There is a community behind them
  - Why reinvent the wheel?
  - Standards facilitate interoperability

# Globus Security

- Control access to shared services
  - Address autonomous management, e.g., different policy in different work-groups
- Support multi-user collaborations
  - Federate through mutually trusted services
  - Local policy authorities rule
- Allow users and application communities to set up dynamic trust domains
  - Personal/VO collection of resources working together based on trust of user/VO



# Virtual Organization (VO) Concept



- VO for each application or workload
- Carve out and configure resources for a particular use and set of users



## GT4 Security

- Public-key-based authentication
- Transport- and message-level authentication (trans. is faster so default)
- Extensible authorization framework based on Web services standards
  - SAML-based authorization callout
  - Integrated policy decision engine
    - > XACML policy language, per-operation policies, pluggable



# Security Tools

- Basic Grid Security Mechanisms
- Certificate Generation Tools
- Certificate Management Tools
  - Getting users “registered” to use a Grid
  - Getting Grid credentials to wherever they’re needed in the system
- Authorization/Access Control Tools
  - Storing and providing access to system-wide authorization information
- Credential management service
  - MyProxy (One time password support)

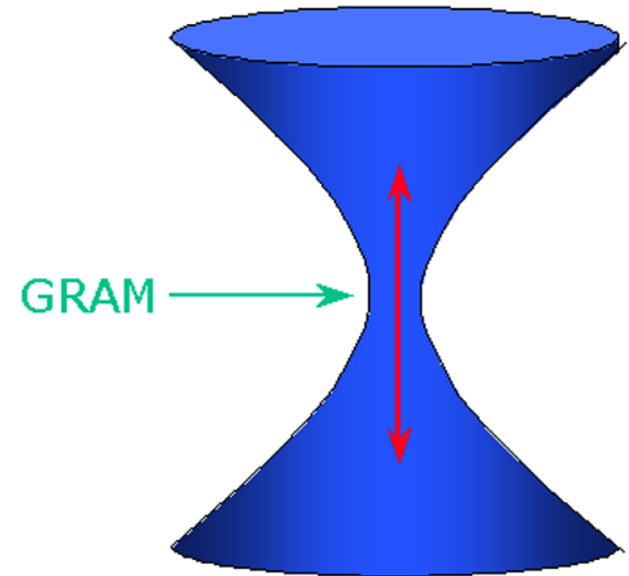


# GRAM - Basic Job Submission and Control Service

- A uniform service interface for remote job submission and control
  - Includes file staging and I/O management
  - Includes reliability features
  - Supports basic Grid security mechanisms
  - Available in Pre-WS and WS
- GRAM is *not* a scheduler.
  - No scheduling
  - No metascheduling/brokering
  - Often used as a front-end to schedulers, and often used to simplify metaschedulers/brokers

## Applications

Metaschedulers, Brokers



Local Management Mechanisms

## GT4 WS GRAM

- 2nd-generation WS implementation optimized for performance, flexibility, stability, scalability
- Streamlined critical path
  - Use only what you need
- Flexible credential management
  - Credential cache & delegation service
- GridFTP & RFT used for data operations
  - Data staging & streaming output
  - Eliminates redundant GASS code



## Execution Management: GridWay

- Ruben will discuss this in a little while



## GT4 Data Management

- **Stage/move** large data to/from nodes
  - GridFTP, Reliable File Transfer (RFT)
  - Alone, and integrated with GRAM
- **Locate** data of interest
  - Replica Location Service (RLS)
- **Replicate** data for performance/reliability
  - Distributed Replication Service (DRS)
- Provide **access** to diverse data sources
  - File systems, parallel file systems, hierarchical storage: GridFTP
  - Databases: OGSA DAI



## GridFTP

- A high-performance, secure, reliable data transfer protocol optimized for high-bandwidth wide-area networks
  - FTP with well-defined extensions
  - Uses basic Grid security (control and data channels)
  - Multiple data channels for parallel transfers
  - Partial file transfers
  - Third-party (direct server-to-server) transfers
  - Reusable data channels
  - Command pipelining
- GGF recommendation GFD.20

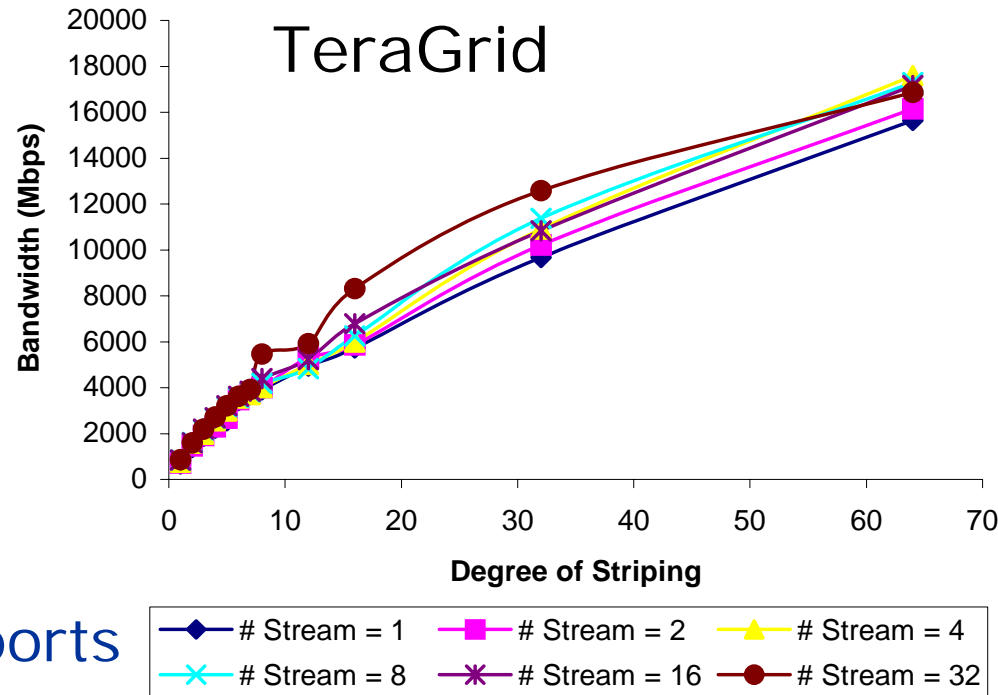




## GridFTP in GT4

- 100% Globus code
  - No licensing issues
  - Stable, extensible
- IPv6 Support
- XIO for different transports
- Striping → multi-Gb/sec wide area transport
- Pluggable
  - Front-end: e.g., future WS control channel
  - Back-end: e.g., HPSS, cluster file systems
  - Transfer: e.g., UDP, NetBLT transport

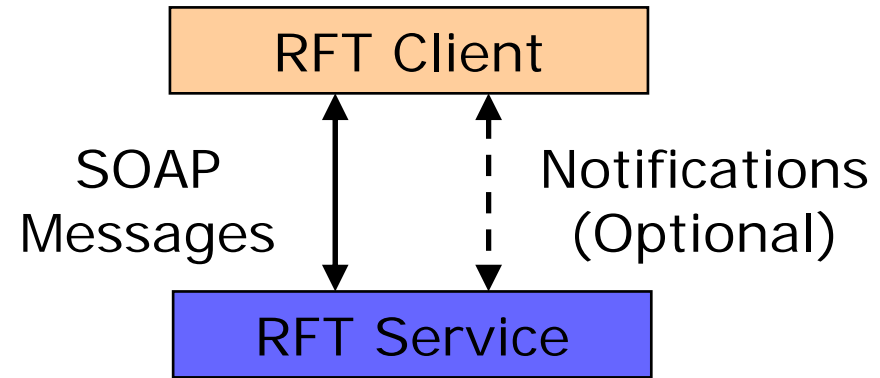
## Disk-to-disk on TeraGrid





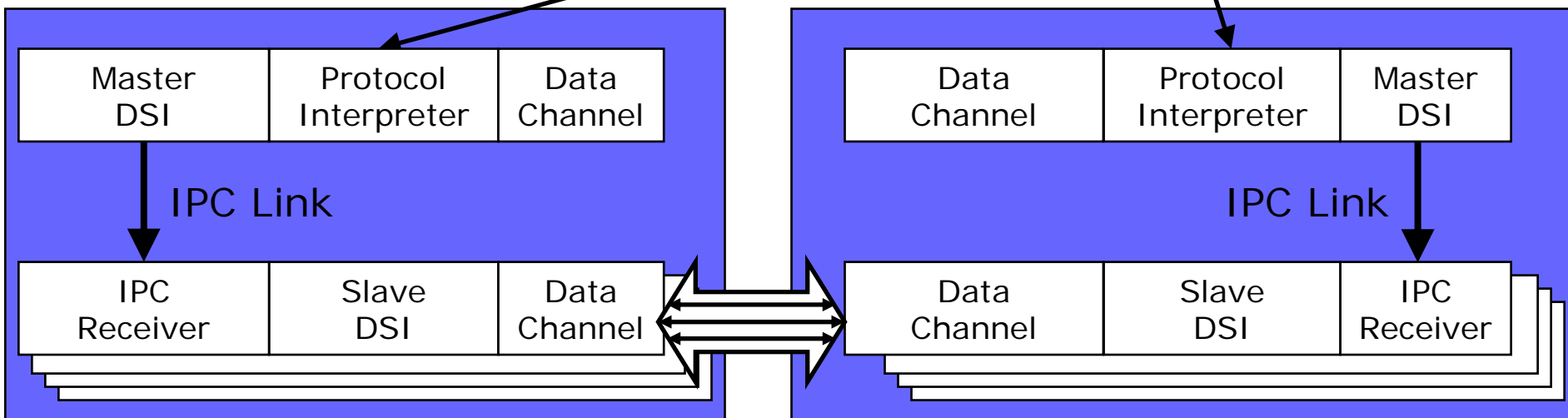
# Reliable File Transfer: Third Party Transfer

- Fire-and-forget transfer
- Web services interface
- Many files & directories
- Integrated failure recovery
- Has transferred 900K files



GridFTP Server

GridFTP Server





## OGSA-DAI



- Web services interface for accessing structured data resources
- Amy Krause is speaking later



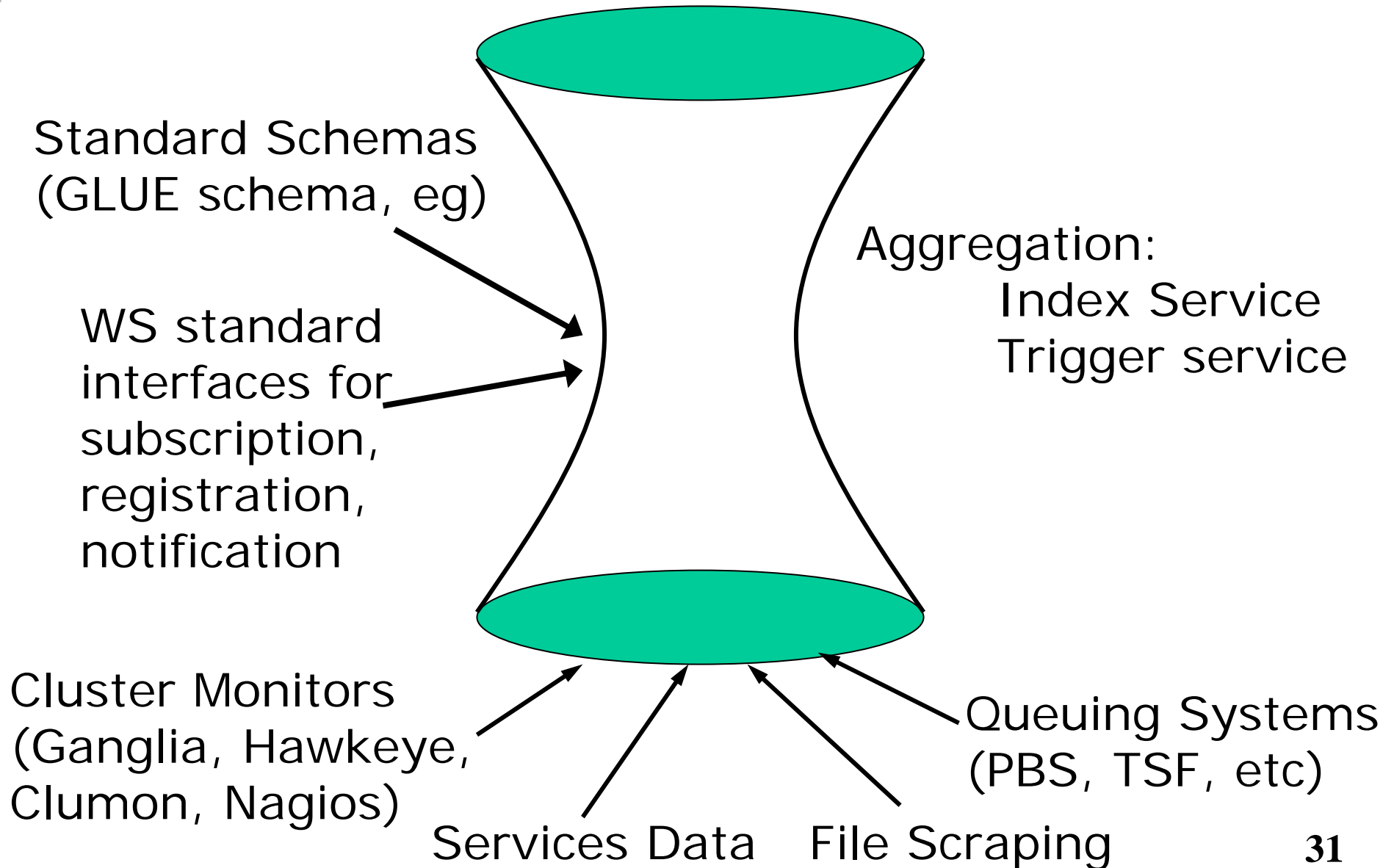
## Monitoring and Discovery System (MDS4)

- Grid-level monitoring system used most often for resource selection
  - Aid user/agent to identify host(s) on which to run an application
- Uses standard interfaces to provide publishing of data, discovery, and data access, including subscription/notification
  - WS-ResourceProperties, WS-BaseNotification, WS-ServiceGroup
- Functions as an hourglass to provide a common interface to lower-level monitoring tools



## Information Users :

Schedulers, Portals, Warning Systems, etc.





## Queue Overview

| Name     | UniqueID | Gram Information |                             |          | LRMS       |         | CPUs  |      | Status  | Jobs  |         |         | Policy Limits   |          |            |              |
|----------|----------|------------------|-----------------------------|----------|------------|---------|-------|------|---------|-------|---------|---------|-----------------|----------|------------|--------------|
|          |          | Version          | Host                        | Port/URL | Type       | Version | Total | Free |         | Total | Running | Waiting | Wall Clock Time | CPU Time | Total Jobs | Running Jobs |
| big      | big      | 4.0.1            | tg-login1.ncsa.teragrid.org | 2019     | PBS-Torque | 2.0.0p7 | 891   | 538  | enabled | 0     | 0       | 0       | 2880            | -1       | -1         | -1           |
| dque     | dque     | 4.0.1            | tg-login1.ncsa.teragrid.org | 2019     | PBS-Torque | 2.0.0p7 | 891   | 538  | enabled | 171   | 50      | 121     | 1440            | -1       | -1         | -1           |
| long     | long     | 4.0.1            | tg-login1.ncsa.teragrid.org | 2019     | PBS-Torque | 2.0.0p7 | 891   | 538  | enabled | 0     | 0       | 0       | 5760            | -1       | -1         | -1           |
| priority | priority | 4.0.1            | tg-login1.ncsa.teragrid.org | 2019     | PBS-Torque | 2.0.0p7 | 891   | 538  | enabled | 0     | 0       | 0       | 1440            | -1       | -1         | -1           |
| debug    | debug    | 4.0.1            | tg-login1.ncsa.teragrid.org | 2019     | PBS-Torque | 2.0.0p7 | 891   | 538  | enabled | 0     | 0       | 0       | 30              | -1       | -1         | -1           |
| quake    | quake    | 4.0.1            | tg-login1.ncsa.teragrid.org | 2019     | PBS-Torque | 2.0.0p7 | 891   | 538  | enabled | 4     | 0       | 4       | 2880            | -1       | -1         | -1           |
| gpfs-wan | gpfs-wan | 4.0.1            | tg-login1.ncsa.teragrid.org | 2019     | PBS-Torque | 2.0.0p7 | 891   | 538  | enabled | 0     | 0       | 0       | 1440            | -1       | -1         | -1           |

## Cluster / Subcluster Overview

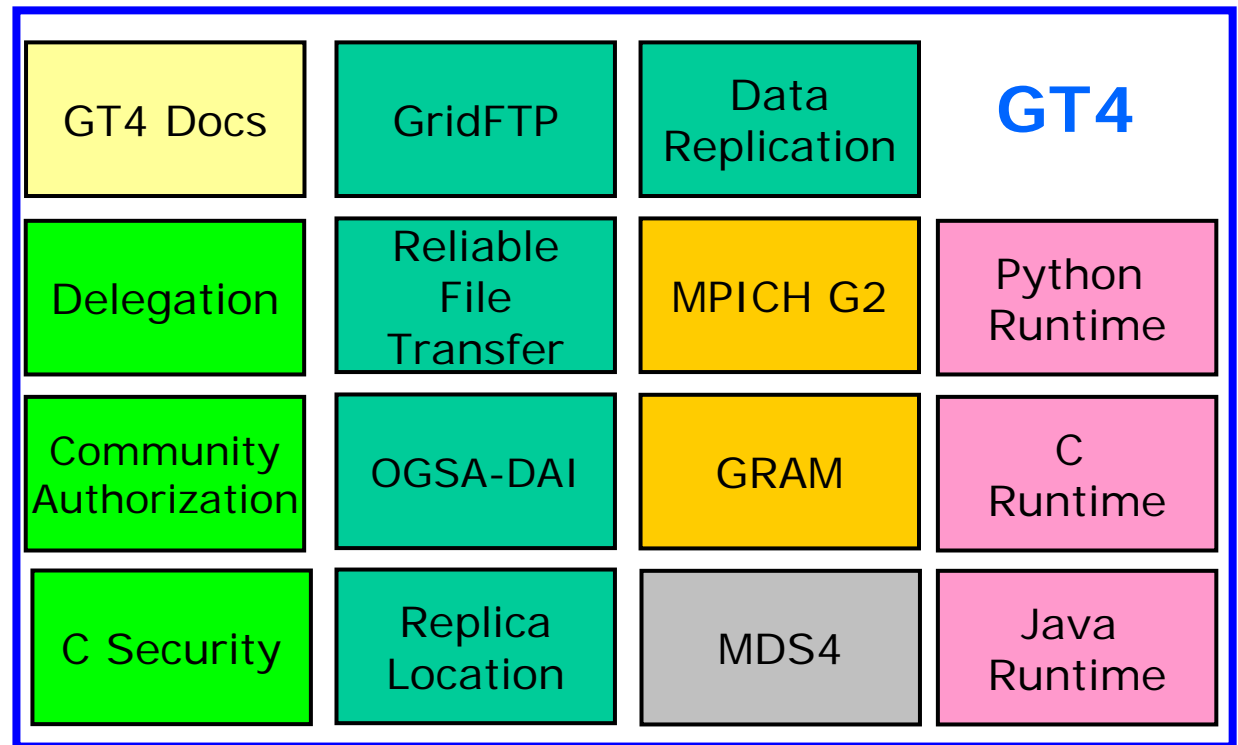
| Type       | Name  | UniqueID                        | Processor |             | Total Memory | Operating System                            | SMP Size | Storage Device |        |                 | TeraGrid Extensions |
|------------|---|---------------------------------|-----------|-------------|--------------|---|----------|----------------|--------|-----------------|---------------------|
|            |   |                                 | Type      | Clock Speed |              |   |          | Name           | Size   | Available Space | Total Nodes         |
| Cluster    | NCSA-TeraGrid                                       | NCSA-TG                         |           |             |              |   |          |                |        |                 | 891                 |
| SubCluster | NCSA-TG-IA64CPU13-FASTIO-HIMEM                      | IA64CPU13-FASTIO-HIMEM          | IA-64     | 1296        | 4061         | Linux2.4.21.SuSE_292.til#1 SMP Fri Jun 3 07 | 2        | entire-system  | 353385 | 91439           | 128                 |
| SubCluster | NCSA-TG-IA64CPU13-FASTIO-LOMEM.ncsa.teragrid.org    | IA64CPU13-FASTIO-LOMEM          | IA-64     | 1296        | 4101         | Linux2.4.21.SuSE_292.til#1 SMP Fri Jun 3 07 | 2        | entire-system  | 353384 | 91435           | 128                 |
| SubCluster | NCSA-TG-IA64CPU15-FASTCPU-GPFSWAN.ncsa.teragrid.org | IA64CPU15-FASTCPU-GPFSWAN       | IA-64     | 1496        | 4106         | Linux2.4.21.SuSE_292.til#1 SMP Fri Jun 3 07 | 2        | entire-system  | 260036 | 10620           | 16                  |
| SubCluster | NCSA-TG-IA64CPU15-FASTCPU.ncsa.teragrid.org         | IA64CPU15-FASTCPU               | IA-64     | 1496        | 4106         | Linux2.4.21.SuSE_292.til#1 SMP Fri Jun 3 07 | 2        | entire-system  | 260036 | 10619           | 615                 |
| SubCluster | NCSA-TG-IA64CPU13-FASTIO-HIMEM-SPARE                | IA64CPU13-FASTIO-HIMEM-SPARE    | IA-64     | 1296        | 4056         | Linux2.4.21.SuSE_292.til#1 SMP Fri Jun 3 07 | 2        | entire-system  | 353372 | 91423           | 1                   |
| SubCluster | NCSA-TG-IA64CPU13-FASTIO-LOMEM-SPARE                | IA64CPU13-FASTIO-LOMEM-SPARE    | IA-64     | 1296        | 4061         | Linux2.4.21.SuSE_292.til#1 SMP Fri Jun 3 07 | 2        | entire-system  | 353385 | 91439           | 1                   |
| SubCluster | NCSA-TG-IA64CPU15-PHASE2-FASTCPU-SPARE2             | IA64CPU15-PHASE2-FASTCPU-SPARE2 | IA-64     | 1496        | 4106         | Linux2.4.21.SuSE_292.til#1 SMP Fri Jun 3 07 | 2        | entire-system  | 260036 | 10620           | 2                   |

## Hosts in Subcluster NCSA-TG-IA64CPU13-FASTIO-HIMEM

| Name                      | UniqueID | TeraGrid Extensions  |
|---------------------------|----------|--|
|                           |          | Node Properties  |
| tg-c001.ncsa.teragrid.org | tg-c001  | all,ia64-compute,compute,ia64-cpu13,fastio,himem,rack40,clos12,stage |
| tg-c002.ncsa.teragrid.org | tg-c002  | all,ia64-compute,compute,ia64-cpu13,fastio,himem,rack40,clos12       |
| tg-c003.ncsa.teragrid.org | tg-c003  | all,ia64-compute,compute,ia64-cpu13,fastio,himem,rack40,clos12       |
| tg-c004.ncsa.teragrid.org | tg-c004  | all,ia64-compute,compute,ia64-cpu13,fastio,himem,rack40,clos12       |
| tg-c005.ncsa.teragrid.org | tg-c005  | all,ia64-compute,compute,ia64-cpu13,fastio,himem,rack40,clos12       |



# Globus Toolkit



Other

Security

Data Mgmt

Execution  
Mgmt

Info  
Services

Common  
Runtime

## GT2 vs GT4

- Pre-WS Globus is in GT4 release
  - Both WS and pre-WS components (ala 2.4.3) are shipped
  - These do NOT interact, but both can run on the same resource independently
- Basic functionality is the same
  - Run a job
  - Transfer a file
  - Monitoring
  - Security
- Code base is completely different



# Why Use Web Service-Based GT4?

- Performance and reliability
  - Literally millions of tests and queries run against GT4 services
- Scalability
  - Many lessons learned from GT2 have been addressed in GT4
- Support
  - This is our active code base, much more attention
- Additional functionality
  - New features are here
  - Additional GRAM interfaces to schedulers, MDS Trigger service, GridFTP protocol interfaces, etc
- Easier to contribute to

# Versioning and Support

- Versioning
  - Evens are production (4.0.x, 4.2.x),
  - Odds are development (4.1.x)
- We support this version and the one previous
  - Currently we're at 4.0.4 (as of 23 Feb. '07) so we support 3.2 and 4.0
  - We've also got the 4.1.1 development release available (26 March '07)



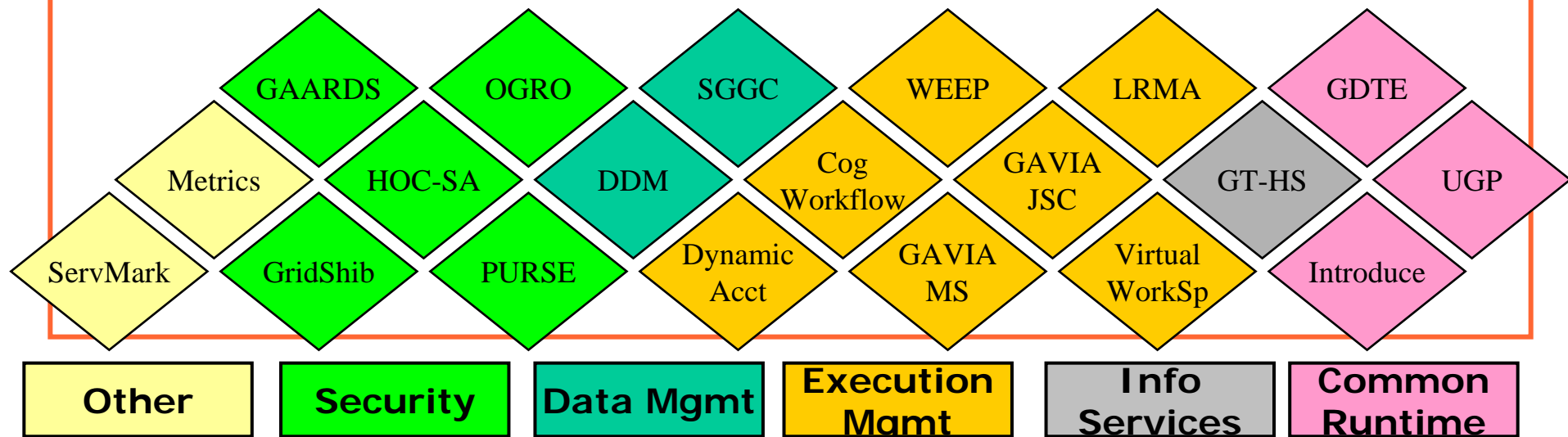
## Several Possible Next Versions

- 4.0.5 – stable release
  - 100% same interfaces, bug fixes only
  - Expected in 1-2 months?
- 4.1.x – development release(s)
  - New functionality
  - Expected every 6-8 weeks (due in about 3)
- 4.2.0 - stable release
  - When 4.1.x has “enough” new functionality, and is stable
  - Underlying q’s with WS-A/WSRF/WS-N spec upgrade
  - Perhaps summer, to be discussed [gt-dev@globus.org](mailto:gt-dev@globus.org)
- 5.0 – substantial code base change
  - With any luck, not for years :)



# Globus Incubator Projects

## Incubator Projects





## Incubator Process in dev.globus

- Entry point for new Globus projects
- Incubator Management Project (IMP)
  - Oversees incubator process from first contact to becoming a Globus project
  - Quarterly reviews of current projects
  - Process being debugged by “Incubator Pioneers”

[http://dev.globus.org/wiki/Incubator/Incubator\\_Process](http://dev.globus.org/wiki/Incubator/Incubator_Process)



## Current Incubator Projects

[dev.globus.org/wiki/Welcome#](http://dev.globus.org/wiki/Welcome#Incubator_Projects)  
[Incubator\\_Projects](http://dev.globus.org/wiki/Welcome#Incubator_Projects)

- Distributed Data Management (DDM)
- Dynamic Accounts
- Gavia-Meta Scheduler
- Gavia- Job Submission Client
- Grid Authentication and Authorization with Reliably Distributed Services (GAARDS)
- Grid Development Tools for Eclipse (GDTE)
- GridShib
- Grid Toolkit Handle System (gt-hs)
- Higher Order Component Service Architecture (HOC-SA)
- Introduce
- Local Resource Manager Adaptors (LRMA)
- Metrics
- MEDICUS
- Open GRid OCSP (Online Certificate Status Protocol)
- Portal-based User Registration Service (PURSe)
- ServMark
- SJTU GridFTP GUI Client (SGGC)
- UCLA Grid Portal Software (UGP)
- WEEP
- Cog Workflow
- Virtual Workspaces



## How Can You Contribute? Create a New Project

- Do you have a project you'd like to contribute?
- Does your software solve a problem you think the Globus community would be interested in?
- Contact [incubator-committers@globus.org](mailto:incubator-committers@globus.org)
- Contact me! I'm at OGF through Thursday  
and happy to help you out  
[jms@mcs.anl.gov](mailto:jms@mcs.anl.gov)



## Contribute to an Existing Project

- Contribute code, documentation, design ideas, and feature requests
- Joining the mailing lists
  - \*-dev, \*-user, \*-announce for each project
  - See the project wiki page at dev.globus.org
- Chime in at any time
- Regular contributors can become committers, with a role in defining project directions

[http://dev.globus.org/wiki/How\\_to\\_contribute](http://dev.globus.org/wiki/How_to_contribute)



## Our Next Steps

- Expanded open source Grid infrastructure
  - Virtualization
  - New services for data management, security, VO management, troubleshooting
  - End-user tools for application development
  - Etc., etc.
- Some infrastructure work
  - How outside projects can join the Toolkit
  - Expanded outreach program ([outreach@globus.org](mailto:outreach@globus.org))
- And of course responding to user requests for other short-term needs

## I should mention some open positions....

- Scientific research programmer, Grid computing research and applications (requisition # 075544)
- Computer systems programmer, grid computing research and applications for workflow (requisition # 075338)
- Software and senior software developers (requisition #075287 and 074800)
- Open Science Grid coordinator of education, outreach and training (requisition # 075427)

<http://jobs.uchicago.edu/>, click "Job Opportunities" and search for requisition number listed



## For More Information

- Jennifer Schopf
  - [jms@mcs.anl.gov](mailto:jms@mcs.anl.gov)
  - <http://www.mcs.anl.gov/~jms>
- Globus Main Website
  - <http://www.globus.org>
- Dev.globus
  - <http://dev.globus.org>